

## Precision Creation of Inter-Gates Insulator

### Inventors

Zhong Dong  
Chuck Jang  
Chunchieh Huang

## Precision Creation of Inter-Gates Insulator

Inventors  
Zhong Dong  
Chuck Jang  
Chunchieh Huang

### 1. Field of Disclosure

[0001] The present disclosure of invention relates generally to semiconductor devices and to the manufacture of the same.

[0002] The disclosure relates more specifically to mass production of insulated gate field effect transistors (IGFET's, MOSFET's) which have plural gate electrodes separated from one another by electrically insulative material. Yet more specifically, it relates to so-called ONO structures which can be used to separate stacked polysilicon gates in electrically re-programmable and nonvolatile memory devices.

### 2a. Cross Reference to Co-owned Applications

[0003] The following copending U.S. patent application is owned by the owner of the present application, and its disclosures is incorporated herein by reference:

[0004] (A) Ser. No. 10/071689 filed February 8, 2003 by Zhong Dong et al and which is originally entitled, "Floating Gate Nitridation".

### 2b. Cross Reference to Patents

[0005] The disclosures of the following U.S. patents are incorporated herein by reference:

(A) U.S. Pat. No. 6,613,695 B2, issued Sept. 2, 2003 to Pomarede et al. and entitled "Surface Preparation Prior to Deposition".

In order to avoid front end clutter, the cross referencing section continues as (2c) at the end of the disclosure, slightly prior to recitation of the

patent claims.

### 3. Description of Related Art

[0006] So-called, FLASH memory devices and alike forms of electrically re-programmable and nonvolatile memory devices have become commercially popular due in part to their ability to store data in a small and dense form factors, their ability to be repeatedly re-programmed, and their ability to retain programmed data in a nonvolatile fashion so that even when power is lost, the stored data is preserved.

[0007] One relatively simple format for such an electrically re-programmable type of nonvolatile memory is known as the vertically-stacked gate configuration. In this configuration, a so-called "floating gate" (FG) electrode is insulatively sandwiched between an underlying, tunnel insulator layer and an overlying, inter-gates insulator (IGI) layer. A so-called "control gate" (CG) electrode is stacked atop the inter-gates insulator (IGI) layer. A channel region with opposed source (S) and drain (D) regions lies under the tunnel insulator layer.

[0008] Because of the way the layers are stacked, the IGI layer (inter-gates insulator layer) is sandwiched between the floating gate electrode (FG) and the control gate electrode (CG). This sandwiching of material layers can create problems as will be further explained shortly. In passing, it is to be observed that other types of more complicated configurations of gate electrodes are possible, including various staggered gate configurations. The simpler, stacked gate configuration will be sufficient for describing the improvement disclosed herein.

[0009] The insulatively-isolated floating gate (FG) of a stacked gate cell is intended to store a certain amount of charge and retain that charge even when external power is turned off. The amount of charge stored on the FG defines the data state of the memory cell. The state of the memory cell can be altered by moving charge into the FG for representing a first data state and by removing charge from the FG for representing another data state. Different

mechanisms may be used for injecting charge into or removing charge from the FG, including hot carrier injection and/or Fowler-Nordheim tunneling. The charged or uncharged state of the floating gate (FG) can be sensed by applying a cell-read voltage  $V_{GS-read}$  to the control gate (CG), where the cell-read voltage  $V_{GS-read}$  is selected to cause a first magnitude of current conduction between the drain (D) and source (S) regions of the cell when the floating gate (FG) is in a first programmed state and to cause no or a different magnitude of  $I_{DS}$  to flow when the floating gate (FG) is in another programmed state. (Some devices store multiple data bits per cell, where each of different amounts of charge trapped within the FG represents a different multi-bit pattern.)

[0010] It is important, for purposes of carrying out the various read and write operations of floating gate type memory cells (e.g., stacked gate cells), to establish an appropriate pattern of electric field intensities across the insulators that surround the charge-storing, floating gate (FG). These electric fields (E-fields) may be established by generating correspondingly appropriate voltages between the control gate (CG), the drain (D), the source (S) and/or substrate (U) regions of the memory cell. Those skilled in the art will appreciate that electric field intensity in dielectric insulators is usually a function of voltage difference (V) divided by dielectric thickness (d) and multiplied by dielectric constant ( $E=kV/d$ ). In order to get consistent results from mass produced devices, it is important to maintain precise control over the dielectric thickness (d) and the dielectric constant (k) of the various insulators which surround the FG of each cell so that same results will occur in one device and the next for a given control gate voltage ( $V_{CG}$ ). Stated otherwise, consistently same capacitive coupling should occur from one mass produced device to the next between the CG, the FG, the source (S), the drain (D) and the substrate (U).

[0011] Persons skilled in the art will appreciate that most insulators are not perfect. A small amount of leakage current can undesirably flow through

them. They can undergo high voltage breakdown. Contaminating chemicals can pass through pinhole or other defects in the insulators. Much progress has been made in the formation of the so-called, thermal gate oxide insulator that lies between the substrate and the floating gate (FG). However, the inter-gates insulator layer (IGI) that is sandwiched between the FG and the CG still suffers from problems.

[0012] The inter-gates insulator layer (IGI) is particularly prone to problems with leakage current and breakdown. Practitioners in the art have developed a so-called ONO solution. According to basic ONO practice, after the material for the floating gate (FG) is deposited (typically it is doped polysilicon), three insulative layers are vapor deposited in sequence. The three insulative layers are respectively composed of silicon-Oxide, silicon-Nitride and silicon-Oxide; hence the ONO acronym. The material for the control gate (CG) is then deposited (typically it is doped polysilicon) on top of the ONO structure. The ONO structure is advantageous because silicon-nitride ( $\text{Si}_3\text{N}_4$ ) generally has a greater dielectric constant ( $k$ ) than does silicon dioxide ( $\text{SiO}_2$ ). The higher dielectric constant of the silicon-Nitride layer in the ONO stack lets cell designers use lower gate voltages than would otherwise be possible for an equal thickness of insulator material made of only silicon-oxide. It is desirable to keep gate voltages relatively small. The ONO approach allows this to happen.

[0013] Despite improvements in vapor deposition techniques, the ONO approach still suffers from the existence of small leakage currents between the FG and the CG. Precise control of ONO thickness and material quality within the ONO stack also remains a problem.

[0014] One improvement over the conventional ONO structure is disclosed in our above-cited U.S. patent application, Ser. No. 10/071,689 ("Floating Gate Nitridation"). Briefly, nitrogen is introduced into the top surface of the doped-polysilicon FG layer after that layer is deposited. The nitrogen concentration is about 1% to about 20% atomic in the top of the FG layer after

nitridation. A higher quality silicon oxide is formed atop the nitridated FG layer, in one embodiment of 10/071,689, by exposing the nitridated, doped-polysilicon to an oxidizing atmosphere. This consumes silicon atoms out of the nitridated region and creates an  $\text{Si}_x\text{O}_y$  layer atop the silicon-depleted-and-nitrogen-containing layer below. The nitrogen-containing layer below acts as an oxidation stop. The technique of 10/071,689 can therefore provide a NONO insulator structure with enhanced insulative properties. Room for yet further improvement is nonetheless available.

#### INTRODUCTORY SUMMARY

[0015] Structures and methods may be provided in accordance with the present disclosure of invention for improving over the above-described art.

[0016] More specifically, in accordance with one set of aspects of the present disclosure, the upper surface of a silicon floating gate (FG) is nitridated to thereby define a non-stoichiometric silicon nitride ( $\text{Si}_x\text{N}_y$ ) surface having a nitrogen content of at least 5% (atomic), and more preferably about 10% to 20% (atomic). This nitridated surface will serve as an oxidation stop. A thin film of intrinsic (essentially undoped) silicon is then deposited on the  $\text{Si}_x\text{N}_y$  surface. Heat and an oxidizing atmosphere are then provided for converting the deposited, intrinsic silicon into a thermally-grown oxide layer ( $\text{Si}_m\text{O}_n$ , where typically  $m=1$  and  $n=2$ ). The non-stoichiometric silicon nitride ( $\text{Si}_x\text{N}_y$ ) layer blocks oxidation below it. Some of the silicon atoms in the  $\text{Si}_x\text{N}_y$  layer may be consumed by the oxidizing atmosphere in the process to create further amounts of thermally generated, silicon oxide ( $\text{Si}_w\text{O}_v$ , where typically  $w=1$  and  $v<2$ ). At least part of the thermally-formed, silicon oxide structure is dopant-free  $\text{SiO}_2$  as a result of the earlier deposition of the intrinsic silicon film.

**[0017]** In one set of embodiments, atomic layer deposition (ALD) is used for precisely controlling the thickness of the intrinsic silicon film. After oxidation of the intrinsic silicon film, a film of stoichiometric silicon nitride ( $\text{Si}_3\text{N}_4$ ) is vapor deposited. Thereafter a silicon oxide film is formed to complete the NONO structure. A control gate (CG) is formed over the NONO structure. Charge leakage between the floating gate (FG) and the control gate (CG) of this embodiment is reduced because the thermally-grown and dopant-free oxide portion of the NONO structure exhibits good insulating properties, particularly because it is grown from intrinsic silicon.

**[0018]** In accordance with an overlapping or somewhat separate aspect of the present disclosure, the lower surface of a control gate (CG) has thermally-grown oxide formed directly under it. In the course of forming an ONO or NONO structure in accordance with this aspect, a thin film of intrinsic silicon (essentially dopant-free) is deposited after the upper, film of stoichiometric silicon nitride ( $\text{Si}_3\text{N}_4$ ) is vapor deposited to define the middle nitride layer of the ONO or NONO structure. The structure is then subjected to an oxidizing atmosphere so as to convert the intrinsic silicon film into a thermally-grown and essentially dopant-free oxide layer. The stoichiometric silicon nitride ( $\text{Si}_3\text{N}_4$ ) layer blocks oxidation below it. In one set of embodiments, atomic layer deposition (ALD) is used for precisely controlling the thickness of the upper intrinsic silicon film. After oxidation of the upper intrinsic silicon film, the material (e.g., doped polysilicon) of the control gate (CG) is vapor deposited. In an overlapped embodiment, both the upper and lower silicon oxide layers of a NONO structure are thermally-grown from intrinsic silicon and both of the intrinsic silicon layers are provided by way of ALD.

**[0019]** An insulating structure in accordance with the disclosure comprises: (a) an oxidation stop layer; and (b) a thermally-grown, intrinsic, silicon oxide layer which has been grown from ALD deposited intrinsic, silicon that had been deposited on said oxidation stop layer.

[0020] An isolation providing method in accordance with the present disclosure comprises: (a) defining an oxidation stop layer above a conductively-doped, first semiconductor layer; (b) providing an intrinsic silicon layer on the oxidation stop layer; (c) oxidizing at least a sublayer portion of the intrinsic silicon layer so as to thereby create a thermally-grown, intrinsic silicon oxide sublayer over the first semiconductor layer; and (d) disposing a conductively-doped, second semiconductor layer above the intrinsic silicon oxide sublayer so that the intrinsic silicon oxide sublayer provides isolation between the first and second conductively-doped, semiconductor layers.

[0021] Other aspects of the disclosure will become apparent from the below detailed description.

#### BRIEF DESCRIPTION OF THE DRAWINGS

[0022] The below detailed description section makes reference to the accompanying drawings, in which:

[0023] FIGURE 1 is a conceptual schematic diagram showing an environment in which it is valuable to form a good, but relatively thin, electrical insulator between two, conductively-doped semiconductor layers (e.g., to thereby form an inter-poly or inter-gates insulator);

[0024] FIGURE 2 is a cross sectional schematic diagram showing how a conventional ONO structure is formed to serve as an inter-poly insulator between stacked gates of a nonvolatile memory device;

[0025] FIGURE 3A is a cross sectional schematic diagram showing a first step (poly-1 deposition) in an IGI fabrication method in accordance with the present disclosure;

[0026] FIGURE 3B is a cross sectional schematic diagram showing a second step (nitridation) in a continuation of the fabrication method introduced by Fig. 3A;

[0027] FIGURE 3C is a cross sectional schematic diagram showing a third step (precise Si deposition) in a continuation of the fabrication method of



Fig. 3B;

[0028] FIGURE 3D is a cross sectional schematic diagram showing a fourth step (beginning of oxidation of precisely deposited and intrinsic Si) in a continuation of the fabrication method of Fig. 3C;

[0029] FIGURE 3E is a cross sectional schematic diagram showing continuation of the oxidation begun by the step of Fig. 3D;

[0030] FIGURE 3F is a cross sectional schematic diagram showing an optional further continuation of the oxidation begun by the step of Fig. 3D;

[0031] FIGURE 3G is a cross sectional schematic diagram showing a fifth step (CVD deposition of SiN) in a continuation of the fabrication method of Figs. 3D-3F;

[0032] FIGURE 3H is a cross sectional schematic diagram showing a set of further steps (CVD deposition of SiO and CVD deposition of CG) in a continuation of the fabrication method of Fig. 3G;

[0033] FIGURE 4A is a cross sectional schematic diagram showing an alternate sixth step (second precise deposition of intrinsic silicon) in a continuation of the fabrication method of Fig. 3G;

[0034] FIGURE 4B is a cross sectional schematic diagram showing an alternate further step (begin oxidation of second precisely deposited Si) in a continuation of the fabrication method of Fig. 4A;

[0035] FIGURE 4C is a cross sectional schematic diagram showing continuation of the oxidation begun by the step of Fig. 4B;

[0036] FIGURE 4D is a cross sectional schematic diagram showing an optional further continuation of the oxidation begun by the step of Fig. 4B;

[0037] FIGURE 4E is a cross sectional schematic diagram showing a yet further step (CVD deposition of doped poly-silicon) in a continuation of the fabrication method of Fig. 4A; and

[0038] FIGURE 4F is a cross sectional schematic diagram showing an alternate structure in which the step of Figs. 3C-3F have been replaced by

CVD deposition of silicon oxide.

#### DETAILED DESCRIPTION

[0039] Figure 1 is a conceptual schematic diagram showing in cross section, an environment 100 in which it may be useful to form a good, but thin, electrical insulator 140 between two, conductively-doped semiconductor layers, 130 and 150. More specifically, the not-to-scale diagram of Fig. 1 shows the structure of a stacked-gate memory cell. The cell 105 is integrally formed as part of a monolithically integrated circuit having a common substrate 110. In one embodiment, the substrate 110 includes doped monocrystalline silicon. Spaced apart source (S) 111 and drain (D) 112 regions may be formed by implant of suitable doping about an oppositely doped, channel portion 115 of the monocrystalline silicon substrate 110. A relatively thin oxide layer (so-called tunnel oxide, or TOX) may be provided on the surface of the channel portion 115. Typically, the TOX is thermally grown from the underlying, monocrystalline silicon 115. (The TOX may be thermally grown prior to threshold implant of dopants into channel portion 115 and prior to the implant doping of the source and drain regions, 111-112. Doping is understood to include the provision of impurity atoms such as boron, arsenic or phosphorus into the semiconductor crystal structure to impart P-type or N-type conductivity to the doped semiconductor region.)

[0040] A first, electrically-conductive gate electrode 130 is formed over the TOX layer 120, typically by chemical vapor deposition (CVD) of doped, polysilicon. A so-called, inter-poly or inter-gates insulator (IPI or IGI) 140 is formed over the first gate electrode 130. A second, electrically-conductive gate electrode 150 is formed over the IGI layer 140. Typically, the second gate electrode 150 is fabricated by CVD of doped polysilicon.

[0041] Patterning of the gate electrodes (130, 150) may occur before or after the source/drain doping implants. Typically, the source/drain doping implants occur after patterning so as to provide for self-alignment with the patterned gates. Post-patterning sidewalls of the first and/or second gate

electrodes 130-150 as well as those of the IGI layer 140 may be surrounded by additional insulative material (e.g., additional silicon dioxide) and/or may be embedded in shallow-trench isolation (STI) wells as may be appropriate. No specific top-view patterning is intended by the conceptual side view of Fig. 1. During device fabrication, electrical connections are typically made to the second gate electrode 150 (hereafter also, the control gate or CG), to the source region (S) 111, to the drain region (D) 112, and to the substrate (U) 110, but not to the first gate electrode 130. The insulation-surrounded first gate electrode 130 is commonly referred to as the floating gate (FG).

[0042] Those skilled in the art will appreciate that the floating gate (FG) 130 is typically used to nonvolatiley store trapped charge 135. This charge 135 may be injected into the FG 130 or removed from the FG 130 by way of hot carrier injection or Fowler-Nordheim tunneling, usually with the injected/extracted moving or tunneling through the tunnel oxide (TOX) 120. When charge is being purposefully injected into the FG 130 or being removed from it, a relatively large, writing voltage ( $V_{GU}$ ) is typically established between the control gate (CG) 150 and the substrate (U) 110. Charge then tunnels or otherwise crosses through the thin TOX layer 120 in the desired direction to add to, or subtract from the amount of charge already present in the floating gate (FG) 130.

[0043] It is desirable to have precise control over the amount of charge being moved into or being removed from the FG 130 during a write cycle because the amount of charge remaining in the FG after a write cycle will determine the memory state of the cell 105. More specifically, the state of the floating gate (FG) will be generally sensed by applying a cell-read voltage  $V_{GS-read}$  between the control gate (CG) and the source region (S) 111 or another reference node. The magnitude of the cell-read voltage  $V_{GS-read}$  will be pre-selected to cause a first magnitude of current,  $I_{DS}$  (not shown) to flow between the drain (D) and source (S) regions of the cell when the floating gate (FG) is in a first programmed state (e.g., representing a binary 1) and to

cause no or a different magnitude of  $I_{DS}$  to flow when the floating gate (FG) is in another programmed state (e.g., representing a binary 0). The charged state of the FG 130 is not limited to just two distinct states. Some memory devices are designed to store multiple data bits per cell, where each of different amounts of charge 135 trapped within the FG represents a different one of the multi-bit patterns (e.g., 00, 01, 10 and 11). The above are just examples. The amount of charge 135 trapped within the FG 130 may alternatively represent an analog value.

[0044] The amount of charge moved into or out of the FG 130 during a write cycle is often determined by a number of factors including the distribution and intensity of electric flux lines through the TOX 120 and the composition and physical dimensions of the TOX. The latter, electric flux lines through the TOX are in their turn, often defined by the voltages present on the control gate (CG) 150, the drain (D) 112, the source (S) 111, and the substrate (U) 110 regions of the memory cell 105. The electric flux lines (not shown) are further defined by the composition and physical dimensions of the IPI/IGI layer 140. More specifically, those skilled in the art will appreciate that electric field intensity in the dielectric insulator or insulators of the IPI/IGI layer 140 will usually be a function of voltage gradient (V) across each of the insulators in the IGI, where the voltages are divided by the respective dielectric thickness (d) and multiplied by the respective dielectric constant (k) of the insulators. This may be roughly expressed as  $E = kV/d$ . (In actual practice, a more detailed analysis is often conducted using Maxwell's equations. There is no need to delve into such details here.)

[0045] It may be understood from the above that mass production fabrication of a multi-gate cells such as 105 shown in Fig. 1 will often call for precise control over the respective dielectric thicknesses (d) and the respective dielectric constants (k) of the various insulators which are provided around the FG 130. Precision is desired so that same results will occur in one device and the next for a given control gate voltage ( $V_{CG}$ ) in combination with

other voltages and further factors (e.g., temperature). In order to get such consistent results in mass produced devices, it is important to maintain consistency with respect to the dimensions and compositions of the insulators that form the IGI layer 140 and the TOX 120. Stated otherwise, consistently same capacitive coupling should occur from one mass produced device to the next between the CG, the FG, the source (S), the drain (D) and the substrate (U).

[0046] Persons skilled in the art will appreciate that most insulators are not perfect. A small amount of leakage current can undesirably flow through them. They can undergo high voltage breakdown. Contaminating chemicals can pass through pinhole or other defects in the insulators. The inter-gates insulator layer (IGI) 140 is particularly prone to problems with leakage current because, unlike the TOX layer 120, the IGI 140 is generally sandwiched between two layers of doped polysilicon: the CG 150 and the FG 130. Artisans have learned that the IGI 140 will not be very good if it made through brute-force oxidation of doped polysilicon in the FG layer 130. Practitioners in the art have developed a so-called ONO solution.

[0047] The conventional ONO solution is illustrated in the cross sectional view (not-to-scale) of Fig. 2. After TOX (220) is grown to a first thickness (Z1) and the doped polysilicon material for the floating gate (FG) 230 is deposited, three insulative layers are deposited in sequence by means of CVD. The three insulative layers are respectively composed of silicon-Oxide (242, having thickness Z2), silicon-Nitride (244, having thickness Z3) and silicon-Oxide (246, having thickness Z4); hence the ONO acronym (240). The material for the control gate (CG) 250 is then deposited (typically it is doped polysilicon) on top of the ONO structure 240. Because stoichiometric silicon-nitride ( $\text{Si}_3\text{N}_4$ ) generally has a greater dielectric constant (k) than does silicon dioxide ( $\text{SiO}_2$ ) or other silicon oxides ( $\text{Si}_x\text{O}_y$ ), the higher dielectric constant of the silicon-Nitride layer 244 in the ONO stack 240 lets cell designers use lower gate voltages ( $V_{\text{GS-read}}$ ,  $V_{\text{GU-write}}$ ) than would otherwise

be necessary for an equal thickness ( $Z2+Z3+Z4$ ) of insulator material made of only silicon-oxide. It is desirable to keep gate voltages relatively small. Thus the ONO solution has become very popular.

[0048] Despite improvements with vapor deposition techniques, the ONO approach still suffers from the presence of leakage current 137 ( $I_{LEAK}$ ) between the FG (130, 230) and the CG (150, 250). Precise control of ONO thickness ( $Z2+Z3+Z4$ ) and material quality within the ONO stack 240 remains a problem.

[0049] One improvement over the conventional ONO structure 240 (Fig. 2) is disclosed in our above-cited U.S. patent application, Ser. No. 10/071689 ("Floating Gate Nitridation"). Briefly, a small amount of nitrogen (e.g., about 1%-20% atomic) is introduced into the top surface of the doped-polysilicon FG layer after that layer is deposited. The nitridated layer is subjected to a high temperature oxygen-containing atmosphere. Silicon atoms in the nitridated layer are scavenged by the hot oxygen-containing atmosphere to form a thermally-grown, silicon oxide. The increased concentration of nitrogen in the silicon-scavenged area stops the oxidation front from moving deeper into the FG layer. A higher quality of silicon oxide may be formed atop the nitridated FG layer by this thermally-growth technique. The technique provides a NONO insulator structure with enhanced insulative properties. Room for yet further improvement is available.

[0050] Fig. 3A is a cross sectional schematic diagram showing a first step 301 (Poly-1 deposition) in an IGI fabrication method in accordance with the present disclosure. Briefly, after TOX layer 320 is thermally-grown to thickness  $Z1'$ , a doped polysilicon layer 330 is deposited to thickness  $Z5$ . In one embodiment, thickness  $Z5$  is in the range of about 1000Å to about 2000Å. TOX thickness  $Z1'$  is in the range of about 10Å to about 100Å, and in a more specific set of embodiments,  $Z1'$  is in the range of about 30Å to about 50Å. The source and drain regions 311, 312 are typically not yet have formed at this stage and are shown in phantom (dashed lines) simply to provide a

point of reference for what will later develop around channel region 315 of the monocrystalline silicon substrate, 310.

[0051] Fig. 3B is a cross sectional schematic diagram showing a second step 302 (nitridation) in a continuation of the fabrication method introduced by Fig. 3A. A nitrogen implant 392 and/or other means are used to introduce nitrogen atoms into top region 335 of the doped polysilicon layer 330. This forms an  $\text{Si}_x\text{N}_y$  composition in region 335 to a depth Z7 below the surface, where y is sufficiently large to provide a concentration of at least 1% atomic nitrogen, and better yet, at least about 5% atomic nitrogen, and more preferably, at least about 10% atomic nitrogen. Various techniques can be used to form an oxidation-stopping gradient wherein nitrogen concentration in region 335 rises to as high as about 20% to 30% atomic nitrogen. Nitridation by way of ion implant may include the application to FG layer 330 (Fig. 3A) of a dose of  $10^{13}$  to  $10^{15}$  atoms/cm<sup>2</sup> or higher of nitrogen at energies in the range 1KeV to about 30KeV to form a homogenous or graded  $\text{Si}_x\text{N}_y$  composition in region 335 to the Z7 depth. Z7 can be as little as about 10Å and should be no more than about 30Å. The ion implant may be followed by a thermal anneal in an inert atmosphere at 850°C to about 1000°C for about 10 to 60 seconds. A Z7 thickness of as little as about 5Å may be obtained if alternate nitridation techniques are used such as DPN (Decoupled Plasma Nitridation) or RPN (Remote Plasma Nitridation) are used. A combination of ion implant and DPN or RPN may be used, followed by optional thermal anneal, to provide a nitrogen concentration gradient where nitrogen density increases moving deeper down the Z7 depth. The relatively high concentration of nitrogen will be used to slow down or stop a descending oxidation front (342 in Fig. 3D) as will be seen below.

[0052] While Fig. 3B shows the surface layer 335 as having just a  $\text{Si}_x\text{N}_y$  composition, it is within the contemplation of the disclosure to incorporate other atoms within this region including an oxygen gradient whose oxygen density decreases moving deeper down the Z7 depth while at the same time

nitrogen concentration increases when moving deeper down the Z7 depth. The oxygen atoms may be ion implanted or otherwise introduced into surface region 335 together with nitrogen atoms. The top of surface region 335 may be treated to enhance surface chemisorption in a subsequent ALD step 303 (Fig. 3C).

[0053] Referring to Fig. 3C, in a next step 303, a layer of intrinsic silicon 341 is deposited to a thickness Z8 on top of the nitridated region 335'. Deposition thickness Z8 can be as small as about 15Å to about 50Å. It can also be thicker if desired. In one embodiment, Atomic Layer Deposition (ALD) is used to form the 15Å-50Å thickness of the essentially dopant-free, silicon layer 341. Other precision forms of deposition 393 of the intrinsic silicon may be alternatively used. Those skilled in the art will appreciate that ALD is related to CVD (chemical vapor deposition) except that only one reactant at a time is pulse-supplied through the deposition chamber. A first of the pulse-supplied reactants (e.g., dichlorosilane  $\text{SiH}_2\text{Cl}_2$ ) adheres to the substrate surface as a chemisorbed monolayer. The second pulse-supplied reactant (e.g., monosilane  $\text{SiH}_4$ ) is provided to react with the chemisorbed monolayer of the first reactant to thereby form a monolayer of the desired reaction product (e.g., silicon). The process of pulse-supplying the first reactant and then the second is repeated N times until a product layer that is N monolayers thick is formed. The thickness of an ALD-defined layer such as 341 can be digitally controlled by the setting of the deposition repetition factor N. Thus, the thickness Z8 of ALD-defined layer 341 can be precisely controlled.

[0054] Referring to Fig. 3D, in a next step 304, an oxygen-containing atmosphere 394 is supplied to the top surface of layer 341 at an appropriate oxidizing temperature, such as about 750°C to about 850°C. The hot oxygen-containing atmosphere 394 converts the exposed, intrinsic Si 341 (Fig. 3C) into thermally-grown  $\text{SiO}_2$  343. Fig. 3D shows the oxidation front 342 as having advanced only partially down into original layer 341 thereby still



leaving some intrinsic Si in region 341'. The thickness, Z8' of the remaining intrinsic-Si 341' is less than the original thickness Z8 of Fig. 3C. Total thickness Z9 of the partially oxidized silicon is greater than Z8 though, because the thickness of thermally-grown oxide is typically twice that of the original silicon material.

[0055] Referring to Fig. 3E, in a next state 305, the oxygen-containing atmosphere 394' continues to be supplied to the top surface of oxidizing layer 343'. Thickness Z9' is about twice that of original thickness Z8 of the original, intrinsic silicon layer 341 (Fig. 3C). When the oxidation front 342' reaches the nitridated region 335', the rate of oxidation slows substantially because silicon nitride acts as a retardant to further oxidation. The degree of retardation will depend on the concentration of nitrogen atoms present as the oxidation front 342' descends down thickness Z7 and continues to bind oxygen to available silicon atoms.

[0056] Referring to Fig. 3F, in an optional next state 306, the oxygen-containing atmosphere 394" has continued to be supplied to the top surface of oxidized upper layer 343". Some of the silicon atoms that were in the  $\text{Si}_x\text{N}_y$  layer 335' of state 305 (Fig. 3E) have been scavenged by the oxidizing process to form a thermally-grown,  $\text{Si}_w\text{O}_v$  layer 344 where the w/v ratio may be greater than 1/2. Underlying region 335" has an increased nitrogen concentration and acts as an oxidation stop. The thickness Z10 of the thermally-grown,  $\text{Si}_w\text{O}_v$  layer 344 may vary and may be in the range of, for example, 0-10Å. Of importance, during the time that region 335" acts as an oxidation stop and the oxygen-containing atmosphere 394" continues to be supplied to the top surface of oxidized layer 343" at an appropriate oxidizing temperature (e.g., about 750°C-850°C), the stoichiometric bonding of oxygen atoms to silicon atoms in dopant-free region 343" continues to become more perfected and as a result, more high quality  $\text{SiO}_2$  material is provided in region 343". Region 343" therefore becomes a high quality insulating region of precise thickness, particularly because it is essentially free

of dopants (P or N type) and because it is thermally-grown and because the thickness of its precursor silicon material (341 of Fig. 3C) was precisely defined by ALD deposition or another, alike precision deposition method. Although ALD is disclosed as a precision deposition method, it is within the contemplation of the disclosure that less precise deposition methods such as CVD may be used to provide the intrinsic silicon 341 on top of nitridated region 335' (Fig. 3C).

[0057] Referring to Fig. 3G, in a next step 307, a silicon nitride layer 345 (e.g.,  $\text{Si}_3\text{N}_4$ ) is deposited (by CVD, ALD or otherwise) to a thickness Z11 on top of the thermally-grown  $\text{SiO}_2$  layer 343". Typically Z11 will be in the range of about 50Å to about 100Å. It is within the contemplation of the disclosure to alternatively use other high-K dielectrics in place of  $\text{Si}_x\text{N}_y$  in layer 345. The good insulative quality of the thermally-grown,  $\text{SiO}_2$  layer 343" reduces the leakage current demands on the overlying, higher-K dielectric layer 345. Examples of alternative higher-K dielectric materials include SiON. The SiON could be alternatively formed by nitridating the top surface of layer 343" and thereby reducing thickness parameter Z9".

[0058] Referring to Fig. 3H, in a next step 308, a silicon oxide layer 346 ( $\text{Si}_x\text{O}_y$ ) is deposited (by CVD, ALD or otherwise) to a thickness Z13 on top of the higher-K dielectric layer 345 (e.g.,  $\text{Si}_x\text{N}_y$ ). This is followed in a subsequent step 309 by deposition (via CVD or otherwise) to a thickness Z15, of the polycrystalline silicon material that will define the conductively doped, CG layer 350. The conductivity-defining dopants of layer 350 may be introduced during the deposition of the polycrystalline silicon material of CG layer 350 and/or afterwards by ion implant.

[0059] Fig. 3H illustrates the completed, multigate structure 360 in which two, conductively doped, polysilicon gate electrodes, 330' (FG) and 350 (CG) are insulated from each other by a NONO structure 340 that includes at least one, thermally-grown  $\text{SiO}_2$  layer 343" that is essentially free of dopants.

This thermally-grown  $\text{SiO}_2$  layer 343" provides enhanced protection against leakage currents. The illustrated NONO structure 340 also includes the nitridated lower region 335", the CVD- $\text{Si}_x\text{N}_y$  layer 345 and the CVD- $\text{Si}_x\text{O}_y$ " layer 346. The thermally-generated,  $\text{Si}_w\text{O}_v$  layer 344 is optional.

[0060] Referring to Fig. 4A, an alternate process path is shown in cross section following the vapor deposition of silicon nitride layer 345 (e.g.,  $\text{Si}_3\text{N}_4$ ) in Fig. 3G. The concentration of nitrogen atoms at the top surface of the  $\text{Si}_x\text{N}_y$  layer 345 of Fig. 4A may be substantially the same as that deeper in the layer, or, if layer 345 is not completely stoichiometric silicon nitride (not all  $\text{Si}_3\text{N}_4$ ) then the concentration of nitrogen atoms at the top surface may be augmented with further nitridation via ion implant, DPN (Decoupled Plasma Nitridation) or RPN (Remote Plasma Nitridation) so that concentration of nitrogen within region 345' is at least about 5% atomic nitrogen, and more preferably, at least about 10% atomic nitrogen. Various techniques can be used to form an oxidation-stopping gradient wherein nitrogen concentration in region 345' rises to as high as about 20% to 30% atomic nitrogen.

[0061] In subsequent step 407, a layer of intrinsic silicon 446 is deposited to a thickness Z12 on top of the nitrogen-containing region 345'. Deposition thickness Z12 can be as small as about 15Å to about 50Å. It can also be thicker if desired. In one embodiment, Atomic Layer Deposition (ALD) is used to form the 15Å-50Å thickness of the essentially dopant-free, silicon layer 446. Other precision forms of deposition 496 of the intrinsic silicon may be alternatively used.

[0062] Referring to Fig. 4B, in a next step 408, an oxygen-containing atmosphere 497 is supplied to the top surface of layer 446 at an appropriate oxidizing temperature, such as about 750°C to about 850°C. The hot oxygen-containing atmosphere 497 converts the exposed, intrinsic Si 446 (Fig. 4A) into thermally-grown  $\text{SiO}_2$  443. Fig. 4B shows the oxidation front 442 as having advanced only partially down into original layer 446 thereby still

leaving some intrinsic Si in region 446'. The thickness, Z12' of the remaining intrinsic-Si 446' is less than the original thickness Z12 of Fig. 4A. Total thickness Z14 of the partially oxidized silicon is greater than Z12 though, because the thickness of thermally-grown oxide is typically twice that of the original silicon material.

[0063] Referring to Fig. 4C, in a next state 409 of the oxidation process 408 started in Fig. 4B, the oxygen-containing atmosphere 497' continues to be supplied to the top surface of oxidizing layer 443'. Thickness Z14' is about twice that of original thickness Z12 of the original, intrinsic silicon layer 446 (Fig. 4A). When the oxidation front 442' reaches the nitrogen-containing region 345', the rate of oxidation slows substantially because silicon nitride acts as a retardant to further oxidation. The degree of retardation will depend on the concentration of nitrogen atoms present as the oxidation front 442' descends down thickness Z11 of the silicon nitride layer and continues to bind oxygen to available silicon atoms.

[0064] Referring to Fig. 4D, in an optional next state 410 of the oxidation process 408 started in Fig. 4B, the oxygen-containing atmosphere 497" has continued to be supplied to the top surface of oxidized upper layer 443". Some of the silicon atoms that were in the  $\text{Si}_x\text{N}_y$  layer 435' of state 409 (Fig. 4C) have been scavenged by the oxidizing process to form a thermally-grown,  $\text{Si}_w\text{O}_v$  layer 444 where the w/v ratio may be greater than 1/2. Underlying region 345" has an increased nitrogen concentration due to the silicon scavenging and acts as an oxidation stop. The thickness Z16 of the thermally-grown,  $\text{Si}_w\text{O}_v$  layer 444 may vary and may be in the range of, for example, 0-10Å, or more specifically if present, about 5Å to about 10Å. Of importance, during the time that region 345" acts as an oxidation stop and the oxygen-containing atmosphere 497" continues to be supplied to the top surface of oxidized layer 443", this being done at an appropriate oxidizing temperature (e.g., about 750°C-850°C), the stoichiometric bonding of oxygen atoms to silicon atoms in dopant-free region 443" continues to become more

perfected and as a result, more high quality  $\text{SiO}_2$  material is provided in region 443". Region 443" therefore becomes a high quality insulating region of precise thickness, particularly because it is essentially free of dopants (P or N type) and because it is thermally-grown and because the thickness of its precursor silicon material (446 of Fig. 4A) was precisely defined by ALD deposition 496 or another, alike precision deposition method. Although ALD is disclosed as a precision deposition method, it is within the contemplation of the disclosure that other deposition methods such as CVD (even if they may be less precise) may be used to provide the intrinsic silicon 446 on top of nitrogen-containing region 345' (Fig. 4A).

[0065] Referring to Fig. 4E, in a subsequent step 411, the polycrystalline silicon material that will define the conductively doped, CG layer 450 is provided (e.g., via CVD or otherwise) to a thickness Z15'. The conductivity-defining dopants of layer 450 may be introduced during the deposition of the polycrystalline silicon material of CG layer 450 and/or afterwards by ion implant.

[0066] Fig. 4E illustrates the completed, multigate structure 460 in which two, conductively doped, polysilicon gate electrodes, 330' (FG) and 450 (CG) are insulated from each other by a NONO structure 440 that includes at least two, thermally-grown  $\text{SiO}_2$  layers, 343" and 443", that are each essentially free of dopants. These thermally-grown  $\text{SiO}_2$  layers, 343" and 443", provide enhanced protection against leakage currents. The illustrated NONO structure 440 also includes the nitridated lower region 335", and the CVD- $\text{Si}_2\text{N}_y$  layer 345". The thermally-generated,  $\text{Si}_w\text{O}_v$  layer 444 is optional.

[0067] Referring to Fig. 4F, an alternate structure 470 is shown wherein the process steps of Figs. 3B-3F have been replaced by CVD deposition of silicon oxide to thereby provide layer 447 sandwiched between the FG electrode 330" and the  $\text{Si}_2\text{N}_y$  layer 345". Source and drain regions 411' and 412' where implanted after patterning of the FG electrode 330". CVD- $\text{SiO}$

layer 447 provides sidewall and top insulation for the FG electrode 330". The overlying parts 345"-443" may be bracketed within a shallow-trench isolation well (having HDP-oxide sidewalls, not shown). The resulting ONO structure 440' within the isolation well includes the thermally-grown and intrinsic, silicon dioxide layer 443". Layer 443" has high quality SiO<sub>2</sub> and therefore reduces charge leakage between the floating gate (FG) 330" and the control gate (CG) 450". Portion 451" is an extension of the control gate which is spaced higher up than Z14 and therefore does not affect electric field concentration near the channel 415' nearly as much as does the descended part of the control gate 450" (the part immediately above SiO<sub>2</sub> thickness Z14).

[0068] The present disclosure is to be taken as illustrative rather than as limiting the scope, nature, or spirit of the subject matter claimed below. Numerous modifications and variations will become apparent to those skilled in the art after studying the disclosure, including use of equivalent functional and/or structural substitutes for elements described herein, use of equivalent functional couplings for couplings described herein, and/or use of equivalent functional steps for steps described herein. Such insubstantial variations are to be considered within the scope of what is contemplated here. Moreover, if plural examples are given for specific means, or steps, and extrapolation between and/or beyond such given examples is obvious in view of the present disclosure, then the disclosure is to be deemed as effectively disclosing and thus covering at least such extrapolations.

#### 2c. Cross Reference to Related Other Publications

[0069] The following publications are cited here for purposes of reference and their disclosures are incorporated by reference:

[0070] (A) U.S. published application No. 2003/0082300 A1 published May 1, 2003, entitled "Improved Process for Deposition of Semiconductor Films", and naming as inventors: Todd et al.

[0071] (B) U.S. published application No. 2003/0166318 A1 published

Sep 4, 2003, entitled "Atomic Layer Deposition of Capacitor Dielectric", and naming as inventors: Zheng et al.

2d. Reservation of Extra-Patent Rights, Resolution of Conflicts, and Interpretation of Terms

[0072] After this disclosure is lawfully published, the owner of the present patent application has no objection to the reproduction by others of textual and graphic materials contained herein provided such reproduction is for the limited purpose of understanding the present disclosure of invention and of thereby promoting the useful arts and sciences. The owner does not however disclaim any other rights that may be lawfully associated with the disclosed materials, including but not limited to, copyrights in any computer program listings or art works or other works provided herein, and to trademark or trade dress rights that may be associated with coined terms or art works provided herein and to other otherwise-protectable subject matter included herein or otherwise derivable herefrom.

[0073] If any disclosures are incorporated herein by reference and such incorporated disclosures conflict in part or whole with the present disclosure, then to the extent of conflict, and/or broader disclosure, and/or broader definition of terms, the present disclosure controls. If such incorporated disclosures conflict in part or whole with one another, then to the extent of conflict, the later-dated disclosure controls.

[0074] Unless expressly stated otherwise herein, ordinary terms have their corresponding ordinary meanings within the respective contexts of their presentations, and ordinary terms of art have their corresponding regular meanings within the relevant technical arts and within the respective contexts of their presentations herein.

[0075] Given the above disclosure of general concepts and specific embodiments, the scope of protection sought is to be defined by the claims appended hereto. The issued claims are not to be taken as limiting Applicant's right to claim disclosed, but not yet literally claimed subject matter by way of